国外信息服务机构的数据管理政策调研与分析

Investigation and Analysis of Data Management Policies of Information Service Organizations Overseas

张闪闪[1,2]　顾立平[1]　盖晓良[1,2]
(1. 中国科学院文献情报中心,北京,100190;　2. 中国科学院大学,北京,100049)

［摘要］　选取国外 10 所大学的图书馆作为调查对象,运用图书馆学信息分析与预测中的跟踪扫描与信息提炼方法,从管理哪些科研数据、如何管理科研数据、如何保障科研数据相关者的权益最大化三个方面分析了科研数据管理的政策要素内容。研究结果表明,在 e-science 环境下,科研数据管理是图书馆未来发展的趋势,我国信息服务机构在制定科研数据管理政策时,可以注重科研数据存储规范(存储义务、存储内容、存储格式)、科研数据质量管理规范(协助创建个人科研数据管理计划、积极开展培训、设计元数据标准、规范命名格式、设计存储时间、建立数据库)、科研数据共享规范(引用要求、使用许可、隐私保护),从累积实践经验完善数据政策以及政策指南。

［关键词］　信息服务机构　科研数据　数据存缴　数据质量　数据共享

［中图分类号］G231　［文献标识码］A　［文章编号］1003-2797(2015)05-0099-11　DOI:10.13366/j.dik.2015.05.099

［Abstract］　By using scanning and information extraction method in information analysis and prediction field, this paper analyzes the contents of research data management policies of the 10 universities' libraries from three aspects——what scientific research data should be managed, how to manage the scientific research data and how to maximize the rights of the stakeholders. The result shows that the research data management has become an inevitable tendency in the e-science environment, and the information service organization should focus on the submitted standard (storage obligation, storage content, storage format), data quality management norm (research data management plan, training, metadata standard, file name, storage time and data repository), and the share policy (quote requirements, license and privacy protection). Also, the information service organization should to improve the data policies and policy guide in the practice.

［Key words］　Information service organization　Research data　Data storage　Data quality　Data sharing

chinaXiv:201606.00015v1

1　科研数据管理离不开信息服务机构的积极实践

科研数据已得到学术界、出版界、图书信息服务界等的高度重视,已有若干学者对此展开讨论[1-5]。科研数据是指在科技活动或通过其他方式所获取的反映客观世界的本质、特征、变化规律等的原始基本数据,以及根据不同科技活动需要,进行系统加工整理的各类数据集[6]。科研数据管理(Research Data Management,以下简称 RDM)包括数据选择、组织、描述、保存、访问、使用和评估等工作过程、实践操作,涵

［基金项目］　本文系中国科学院项目"中国科学院文献出版领域择优支持人才"(院 1434)研究成果之一,部分内容曾在 2014 年学科馆员服务学术研讨会上发表。

［作者简介］　张闪闪,女,博士研究生,研究方向:数字图书馆建设、知识组织、文献计量;顾立平,男,副研究员,博士,研究方向:科研数据权益管理、开放获取政策、知识产权与著作权,Email:gulp@mail.las.ac.cn;盖晓良,男,硕士研究生,研究方向:科技政策与管理情报、信息资源管理。

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

盖了从制定标准、数据映射至不同仓储到元数据创建、语义注释和文献链接等广泛的活动，处理后的数据具有增值的普遍特征[7]。这多方面的 RDM 强调与外界环境(政策、机构)以及人之间的协调[8]。

目前，国内外已有若干研究，国外如 Kim[9] 指出 RDM 服务可以增加科研人员对自己数据的管理，Shen 和 Virgil[10] 调研了美国约翰霍普金斯大学的数据知识库的实施情况，Stephen Kutay[11] 分析了科研教育机构在进行数据资源管理时需要注意的一些问题，Davidson[12] 分析了目前英国正在实施的数据管理项目，著名信息搜寻行为研究者 Gary Marchionin 在武汉大学介绍了 iSchool 的 RDM 经验[13]，Jelte[14] 也曾说数据共享应该是规则而不应该是特例；国内的研究主要集中于国外实践管理政策的归纳总结，如顾立平[15] 分析了 RDM 中的权益问题，吴振新[16] 对比了 5 所大学 RDM 的政策，吴新年[17]、沈梦轩[18] 分析了 RDM 政策的服务和实践，黄如花从科学元数据的功能需求分析、标准建设等方面分析了国内外科研数据元数据的研究进展[19]，杨云秀、顾立平等人也对国外科研教育机构和科研资助机构的数据政策进行了分析[20-21]。研究美国、欧盟等国外知名信息服务机构的 RDM 政策有助于我国信息服务机构制定更加完善的数据管理政策。

近年来，国内已有引领 RDM 服务的实践，比如 2011～2012 年，在 CALIS 三期项目建设中，武汉大学图书馆尝试开展了 RDM 服务：确定试点院系，搭建共享平台，并开始面向全校提供数据存储和共享服务[22-25] 机制和平台。所取得的几项研究成果包括：对比国外高校的科研数据生命周期的管理模型[26]、国内外高校数据共享平台的整理[27]、国外政策以及国内科研教育体系的科研数据服务发展[28]、国内高校用户对 RDM 的需求调查[29]、武汉大学图书馆在社会 RDM 方面的实践[22]、武汉大学图书馆实行 RDM 的案例分析[23]，以及建立科研数据知识库的经验总结[30] 等。由武汉大学、中国科学技术信息研究所、澳大利亚南昆士兰大学等于 2014 年 10 月 26～27 日联合举办的 COINFO 2014 专门介绍了科研数据的引用、发布与共享[31]，这些具体实践都为我国信息服务机构开展 RDM 提供了良好的基础。

RDM 是解决科研成果完整性、科研数据规范化问题的有效方式，通过对科研数据的管理，能够提高研究的可见性，促进学术成果的严谨性，提高研究的质量，通过再利用提高资源的效率，促进科研人员发现更多更新更有价值的问题，提高科学的完整性以及推动科学的进一步发展[32]。

中国科学院文献情报中心对 RDM 开展了相关的研究和实践。在此基础上，为积极面对国内外的数据管理政策变化，结合现阶段科研人员的实际需求，本文从信息服务实践的角度，提炼了 DRM 的政策要素，以期为我国的 RDM 政策制定提供参考借鉴。

2　信息服务机构 RDM 政策要素的框架分析

科研数据的管理更加注重完整流程，从数据创建、存储，到后续的管理和完善，直至实现科研数据的共享，都是 RDM 的内容，如图 1 所示。

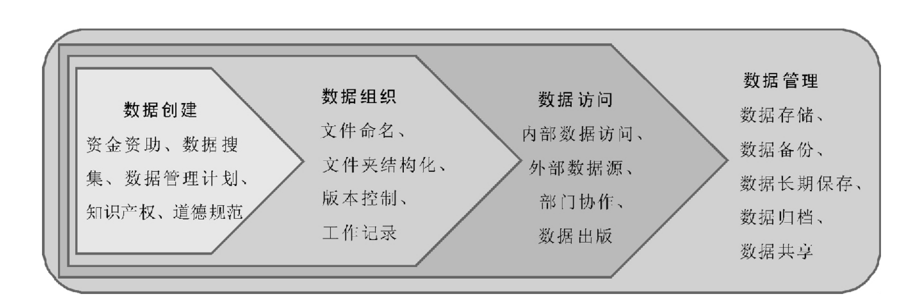


图 1　RDM 内容框架[33]

在遵循科研资助机构与科研教育机构的数据管理政策的基础上，制定良好的 RDM 政策，人们通常考虑“机构要管理什么数据，数据创建之后如何进行存储和保管，才能在保证相关人员权益的同时实现数据

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

的最大化利用”，如果按照这个思路对信息服务机构的科研数据政策进行分析讨论，则可进行研究提问：

提问1：管理什么？信息服务机构要管理哪些科研数据？需要遵循怎样的规则？

提问2：如何管理？科研数据创建之后，信息服务机构要如何实现对这些数据的保管、整理？

提问3：怎么使用？信息服务机构要如何保障各方面的权益，从而实现数据的最大化利用？

根据这三个研究提问，借鉴《前瞻导论》[34]的政策内容分析方法，细分为若干可观测的研究问题，如表1所示。

信息服务机构是负责收集、处理、加工信息的机构，也是对科研数据实施管理的机构，主要包括科技情报所、图书馆、信息中心、咨询公司等，无论是隶属于学校的图书馆，还是专门的信息服务中心，其科学研究一般都受到科研资助机构的资助，所以信息服务机构的政策既要遵从科研资助机构关于RDM的政策要求，又要有自身独特的数据政策规定，部分信息服务机构的科研数据政策还要符合所在大学的政策要求。

以科技信息政策中心编译的《欧盟图书馆实施科研数据管理的案例汇编》为基础，实施滚雪球法，选取

表1　信息服务机构的RDM政策的分析框架

研究提问	观测问题
管什么？	1.RDM为什么由信息服务机构来执行，而不是其他部门？ 2.科研数据存储的数据内容有无明确规定？ 3.科研数据存储的数据格式有无明确规定？
如何管？	1.是否要求科研人员提交DMP？DMP都包括哪些内容？ 2.信息服务机构是否提供RDM的培训与指导？ 3.科研数据的结构化和规范化格式是怎样的？ 4.数据文件的归类需要遵循什么样的命名规则？ 5.科研数据的存储时间有无明确的相关规定？ 6.科研数据的存储位置有无明确规定？对数据库有无要求？
怎么用？	1.科研数据开放共享中，如何保证数据产生者的权利？ 2.科研数据使用许可有无明确规定？ 3.科研数据的隐私保护是如何规定的？

具有代表性和借鉴意义的10个研究对象，如图2所示。以下简称剑桥大学图书馆(LCAM)、麻省理工学院图书馆(LMIT)、耶鲁大学图书馆(LYL)、纽约大学图书馆(LNY)、斯坦福大学图书馆(LST)、约翰霍普金斯大学图书馆(LJH)、康奈尔大学图书馆(LC)、伦敦国王学院图书馆(LKCL)、维也纳大学图书馆(LU)、内梅亨大学图书馆(LR)。

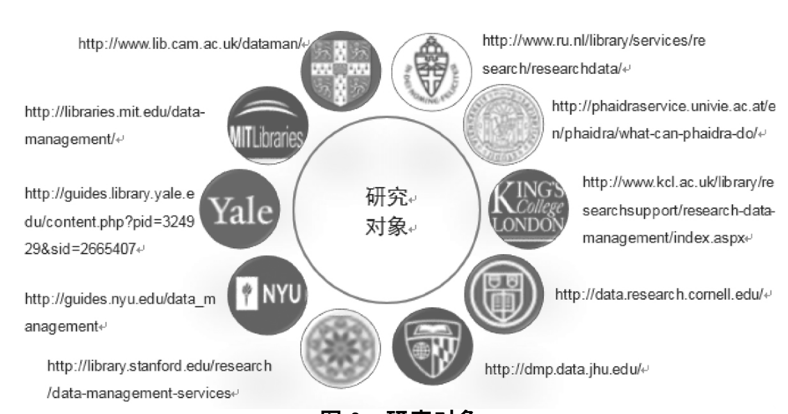


图2　研究对象

3　信息服务机构的科研数据存缴内容要点

3.1　信息服务机构支持RDM的多重角色

在E-science环境下，科研数据呈现指数型增长趋势[35]。图书馆能够从不同角度提供各种数据服务，Auckland[36-37]指出图书馆员能够提供数据咨询服务、数据分析、数据存储等服务；Anna也认为图书馆能够在国家数据管理规划和发展策略制定、高校RDM实施、RDM人才培养方面发挥重要作用[38-39]；Cox[40]提出图书馆在RDM中扮演着政策和宣传、支撑和培训、审核和数据库存储的角色；Luis提出图书馆员在未来RDM中肩负着“数据监管员(data curators)”的责任[41]；Lyon将图书馆的潜在作用与科研生命周期模型进行映射，指出图书馆能够在RDM中提供RDM收

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

集、RDM 策划、RDM 信息服务、科研数据引用、RDM 培训、科研数据许可、科研数据评估、科研数据存储、科研数据获取和科研数据影响力十个层面上的服务[42-43]。也正是由于科研数据具有的科研价值、经济价值和社会价值,并且是独一无二的,一旦遭到破坏将无法恢复,而信息服务机构作为信息的管理中心,既能够对数据资源进行标引、分类及整理,又能确保数据资源的安全性,因此信息服务机构在科研数据的保存和利用方面具有不可替代的作用。

3.2 信息服务机构要求作者提交的科研数据内容

科研数据内容是反映客观事物基本特征及其变化规律的原始基础数据。内容包括但不限于:①调查数据;②实验数据;③访谈数据;④教育数据;⑤工作文件等。许多大学图书馆在进行 RDM 时都列举了所要存储的科研数据的具体类型,比如 LR 存储的科研数据包括收集到的各类数据(如访谈数据、视频数据、调查数据、定量数据、核磁共振数据、观测数据、实验数据等)[44]。LYL 也给出存储的数据包括各种调查数据、模拟数据、实验数据和其他可编译数据等[45]。LMIT 管理的数据内容包括会议论文、图片、预印本、技术报告、论文、工作文件、科研数据集等[46]。LJH 也规定数据可以是实验数据、观察数据等,但是数据内容必须是包括可以进一步使用和评价的记录、已发表的科研成果等,无法再进一步利用的数据是不能被存储的[47]。也有一些机构比较特殊,对数据存储内容没有特定要求,如 LU 存储的科研数据形式可以是观测数据、实验数据、也可以是派生或衍生数据、引用或规范数据以及生成数据,除此之外,还可以有教学意义上的科研数据[48]。

3.3 信息服务机构对科研数据存储格式有所要求

科研数据的格式影响着其开放和引用,是确保科研数据能够长期获取的关键,主要包括:①开放的;②不依赖于软件的;③非专业;④未加密;⑤非编译格式等,这种格式的数据更容易被保存和使用,也不会受到软件限制。如 LMIT 指出为了保证科研数据能够在未来一段时间内可以被发现,提出了数据保存时的优选格式——ODF、ASCII、MPEG-4、TIFF 或 JPEG2000 以及 XML[49]。LKCL 也提出无论在分析时使用何种格式,但是在进行数据保存时最好使用通用的格式,尽量不要存储诸如 RTF、Excel、SPSS 这种依赖于软件的文件格式[50]。LST 虽然也要求存储开放格式的数据,但是如果格式转换会造成数据丢失,则也可以存储专业格式及其 readme. txt 文档和相关读取软件[51]。也有部分图书馆比较例外,如 LYL 认为任何类型的数字文件(录像、文本、照片、数字等)都可以被存储[52]。

4 信息服务机构科研数据的质量管理要点

4.1 协助创建个人 DMP

良好的数据管理计划(Data Management Plan,简称 DMP)是科研数据共享的前提,它不仅可以帮助科研人员管理科研数据,满足资助者的要求,而且还能够提高资源的共享度。因此,很多机构都提出通过 DMP 来帮助科研人员实现数据资源的管理。DMP 的内容一般包括:①项目、实验、数据的描述;②文档、组织和存储;③访问、共享和重用;④归档等[53](见图 3),如 LMIT、LST、LCAM、LYL、LNY、LC、LJH、LR 等均提出需要清楚地描述科研数据的创建、内容、管理人员、发布、共享权利、存储位置等相关方面的政策,编制成在线模板供科研人员使用,且提供了科研数据创建工具 DMPTool,帮助科研数据成为现在和未来可用的最佳实践[53-59]。此外 LMIT、LC 等还提供了 DMP 的实例链接。

4.2 积极开展 RDM 的培训工作

图书馆员是科研管理工作的主要执行者,其对科研数据的认识程度直接决定了 RDM 工作执行的好坏,因此,很多学校专门开设了培训课程,主要包括:①针对于图书馆员的培训,使其对 RDM 工作所需的专业技能有更全面的了解,为科研人员提供管理科研数据的支持;②针对科研人员及学生的培训,帮助他们了解科研数据的重要性,增强对科研数据的认识。LMIT[60-61]、LCAM[62]、LNY[63]、LJH[64] 等提供了各种各样的数据管理培训班帮助科研人员学习如何管理科研数据、如何进行科研数据的文件组织以及如何进行科研数据的版本控制等课程,所有课程都可以通过主页进行查找。LKCL 也开设了类似的课程,除了提供相关培训材料和课程时间安排外,还设置了课程反馈意见调查,内容主要包括[65]:①你希望从课程中学到

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

什么？②你已经学到了什么？③你还想学习哪些内容？④你认为课程的优点是什么？⑤你认为课程还有哪些地方需要调整？⑥你愿意向其他人员推荐该课程吗？⑦其他建议。LR 建立了一个专业科研数据中心,要求所有员工都必须拥有 RDM 服务(开发)的知识,还专门为博士生开设了 RDM 课程[66]。

项目、实验、数据的描述
1. 这项研究的目的是什么？
2. 数据是什么？应该如何收集？
3. 这项研究将产生多少数据？
4. 收集数据需花费多少时间？多久会发生改变？
5. 如何使用他人数据？数据的获取途径是什么？
6. 谁负责管理数据？如何确保DMP实施？

文档、组织和存储
1. 创建何种文档会便于研究人员理解？
2. 元数据是如何管理和元存储的？
3. 文件格式是怎样的？是否符合开放标准？
4. 使用什么类型的文件格式？
5. 目录和文件命名有什么限制？
6. 本地存储和备份程序是什么？
7. 阅读或者查看这些数据的工具或软件有哪些？

麻省理工学院图书馆的个人DMP

归档
1. 数据如何归档？数据将会被存储在归档库或知识库中以供长期访问吗？
2. 需要专业知识库吗？如果不需要，是否可以考虑存储在DSpace@MIT？
3. 数据需要匿名或转换为更稳定的文件格式吗？
4. 使用这些数据需要软件或工具吗？软件或工具也会一并存档吗？
5. 数据保留多长时间？3-5年、10年、还是永远？

访问、共享和重用
1. 谁有权利管理这些数据？
2. 哪些数据将被共享？何时？如何共享？
3. 数据共享会产生隐私、伦理或保密问题吗？
4. 数据和其他项目信息的知识产权归谁所有？
5. 资助项目内数据共享有无专利或技术许可？
6. 如果在杂志上发表，需要底层数据吗？
7. 数据有时滞期吗？
8. 是否允许重用、再分配、或创建新工具？

图 3　麻省理工学院图书馆的个人 DMP[53]

4.3　设计符合需求的科研数据元数据标准

科研数据元数据标准包括但不限于[67]:①标题;②创建者;③标识符;④日期;⑤主题;⑥资助者。由于不同学科特点不同,因此各信息服务机构也都针对自身特点给出了不同的元数据标准。LMIT 规定了科研用户上传数据时的元数据内容(见表 2),并且建议用户在搜集和创建元数据时应该对数据产生、组织以及加工过程、缩写等都做好笔记。

LYL[69]、LNY[70]、LST[71]等也都给出了科研数据上传的元数据描述标准,指出采取何种标准依赖于数据的类型、规模以及所属学科,而 LC 还特别提供这种说明性文件(命名为 readme)[72-73]。相比科研用户自身存储数据,由图书馆统一管理下的科研数据资源更加结构化,有利于储存、检索和挖掘分析等相关利用。

4.4　规范化科研数据的文件命名格式

文件命名格式可以反映文件的特征、内容,因而

表 2　LMIT 规定上传的科研数据元数据内容[68]

元素	描述
标题	数据集或科研项目的名称
创建者	创建人/创建组织的名称和地址
标识符	用于识别数据的号码,即使它只是一个内部项目参考号码
日期	与数据相关的关键日期(包括项目起始日期、数据修改日期、发布日期、时间段)
主题	数据主题或内容的主要关键词、短语
资助者	科研项目的资助机构
权利	数据的知识产权
语言	可使用的语言类型
位置	数据的物理位置、空间信息
方法	数据是如何产生的(包括所使用的软硬件设备、实验协议及其他内容)

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

需要有足够的描述信息以识别数据文件。文件命名包括但不限于：①项目缩写；②实验名称；③实验位置；④观察员；⑤观察年份；⑥数据类型；⑦版本号；⑧文件类型，同时为了便于自动化处理，文件名中一般很少涉及特殊字符以及空格，通常会使用“—”或“_”来代替，比如LMIT[74]的其中一种文件命名格式为：

[type of file]_[specimen number]_[version]_[collector]_[YYYYMMDD]_[geolocation].[ext]

比如蝴蝶项目中一张图片的名称即为 image_12345_v1_mcneill_20140117_42.3598N71.0921W.tif

从这个名称中我们可以清楚地看出图片所描述的内容，这对于数据文件的归类整理、资源共享、查找都有非常重要的作用。LYL[69]、LC[75]、LST[76]等也给出了类似的数据文件命名格式。

4.5　协助设计合理的科研数据存储时间

数据的存储时间一定程度上决定了数据的发现和访问服务，对支持科研数据的生命周期具有非常显著的意义。存储时间一般包括：①有时间限制(4年、7年、10年)；②无时间限制(永久保存)。如LKCL规定，如果是由项目资助的已发表科研成果，那么将会保存7年，而未受到项目资助的，会根据情况保存4年；对于没有历史价值且可以二次使用的科研数据，根据兴趣保存4年，而对于有历史价值的科研数据，则在存储时要识别主题和国家，具体时间没有限制[77]。也有图书馆对所有数据采取同样的保存期限，如LR要求科研人员致力于保持与他们的学术出版物相关的数据可获取的时限至少为10年[44]；LU开发了一个系统——Phaidra，倡导持续引用的数据政策：“一旦在Phaidra中，就一直在Phaidra中”，以此来实现数字资源和资产的永久托管、存储和索引[78]。

4.6　建立数据知识库，提供外部链接，协助科研用户存储科研数据

数据知识库的使用可以帮助科研人员管理数据、统一引用标识符、提高数据的可发现性以及促进数据的长期保存。在数据的存储位置上，通常有：①学校的数据知识库；②公共数据知识库，比如LJH、LCAM、LU、LST都分别开发了自己的知识库JHU Data Archive[79]、DSpace平台[80]、u:scholar[81]、SDR[82]，鼓励师生存储论文、科研数据等资源，但是LCAM还建议关键性的数据应该存储在不同位置而不仅仅是一个数据库中[83]。LMIT还对数据类型进行了区分，其中由LMIT和HP公司合作研发的DSpace平台、麻省理工与哈佛数据中心共同研发的DVN平台，就是专门储存学校教师、研究人员、部门、实验室人员的数字科研成果、教学成果的数据知识库；对于社会科学，可以存储在ICPSR上；此外，图书馆还可以协助科研人员将数据存储在DataBib、re3data.org、OAD上[84]。LC在NSF的资助下，建立了康奈尔地理空间信息库(GUGIR)和美国农业部经济、统计与市场信息系统(USDA-ESMIS)两个数据储存库，供研究人员存储、分享和发布数据，在此基础上，还开发了数据阶段型存储库(Data staging repository, DataStaR)[85]。LNY[86]、LYL[69]等虽然没有专门的知识库，但都提供了不同学科知识库的链接供用户参考。

5　信息服务机构的科研数据开放共享的要点

5.1　规范化科研数据的引用格式

数据引用不仅指文献之间的参考文献，也可以指数据和文献、数据和数据集、数据和数据之间的多重关系[87]，这对于提高科研人员的个人绩效具有一定的影响。在引用科研数据时，一般包括：①作者；②标题；③出版年；④出版商/经销商；⑤URL、标识或其他访问位置等。目前，数据知识库往往赋予每个数据集一个唯一标识符(DOI)，这个标识符能够快速准确识别数据，对于数据引用至关重要，部分机构库还提供数据引用格式，供用户参考。比如LMIT[88]、LJH[89]的社会科研数据服务(Social Science Data Services)指出了数据引用的意义，并规定引用数据时必须包含作者、标题、出版年、出版商/经销商和URL、标识或其他访问的位置五项内容，给出了数据引用的案例及其链接，LYL[90]、LC[91]等也给出了同样的数据引用格式，科研人员可以参考给出的格式实现对数据的引用。LCAM以FAQ的方式介绍了数据的引用，他们认为数据引用没有统一格式，不同学科数据的引用格式也不同，同时给出了密歇根州立大学图书馆以及LMIT数据引用格式标准的链接，供用户参考[92]。

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net



5.2 协助制定合理的科研数据使用许可及使用范围

知识成果的使用权利必须通过一系列的条件与限制予以界定[93-94]。

在使用许可上，可以是：①Creative Commons；②Open Data Commons。如 LCAM 鼓励用户使用 DSpace 平台上的数据，为了鼓励对平台上资源的再利用，规定了 Incremental 和 DataTrain 项目上的数据许可在 BY-NC-SA-2.0，U.K 之下，PrePARe 项目上的数据许可在 BY-SA-2.0 之下[95]。对在 SDR 数据库中数据的使用许可上，LST 给科研人员提供了两种选择——Creative Commons（以 CC0 为主，其他 CC BY 为辅）和 Open Data Commons（PDDL、ODC Attribution License、ODbL），科研人员可以根据自己的意愿来选择使用哪种许可[96]。LMIT 将数据分为自己创造/搜集的数据和从其他来源搜集的数据，其中自己创造的数据应该遵从 CC 许可的规定，而其他来源数据只能用于教育或科研使用，而不能作他用[97]。

在使用范围上，可以是：①机构内部人员使用；②机构以外人员使用。如 LKCL、LR 等都明确规定，作者提交数据时，应该标识出谁可以获取及使用这些数据？数据将如何使用的问题。LR 还规定所有发表的科研数据都可以被反复检验，数据资源是对图书馆内所有成员单位（用户、信息服务机构）共享的，当然前提是必须谨慎处理科研数据和诚信学术实践之间的关系[44]。LMIT 的 DSpace 平台和 DVN 平台中的科研数据是面向全世界用户开放的，最大限度地实现了开放共享。

5.3 合理划分科研数据类型，注重科研数据的隐私保护，创造良好的保存环境

安全的数据保存环境是保证科研数据有效传递的重要基础。当涉及机密数据、敏感数据或受试者数据时，科研数据的管理就变得更加重要了。数据类型一般包括：①禁止类数据；②限制类数据；③保密类数据；④公开类数据。数据的隐私保护一般根据数据的类型进行区分，如 LST 就把数据分为了以上 4 类并提供了三种数据存储方式，Secure AFS——四种类型均可存储，但其中的 Standard AFS 只允许存储公开和保密型数据；Secure IGFS——存储公开、保密和限制型数据，Standard IGFS 允许存储公开和保密型数据；Online Archive Storage——允许存储公开、保密和限制型数据[98]。在数据共享上，要对数据类型实施修改以保证数据能被共享，同时鼓励科研人员保留多个版本的科研数据，一种是适合公开的，另一种是在严格限制条件下可以展开进一步研究的[99]。当涉及到人类受试者数据时，共享将会受到斯坦福机构审查委员会（IRB）的监督，而且对科研分析工具也有要求，在搜集和管理病人健康或其他敏感类信息时，只能使用 REDCap、STRIDE、Qualtrics 三种工具，Excel 不能充当搜集工具，只能分析不含个人信息或匿名化的数据[100]。LKCL 更是规定，个人数据的开放使用权利由个人来决定，即便是学院，也必须是在非常必要的情况下才使用个人科研数据，并且匿名数据是首选数据源[101]。

6 结论与展望

6.1 信息服务机构的政策要素

综上所述，针对本研究所提出的三个研究问题可将信息服务机构的 RDM 政策要求归纳为三方面：科研数据存缴内容规范、科研数据质量管理规范、科研数据共享规范，如表 3 所示。

信息服务机构作为 RDM 的最佳实践者，对科研数据的筛选、DMP 的制定、科研数据的描述与保存、科研数据的共享与使用都有重要的指导作用，因此，科研信息服务机构在制定 RDM 政策时，可以增加与数据存缴、数据质量、数据共享相关的政策内容。此外，在政策制定上，信息服务机构往往还会受到科研资助机构和科研教育机构的指导，因此需要遵循两者在 RDM 方面的政策。

6.2 科研数据政策的制定需要从实践中来，并接受实践检验

本文通过文献综述，采取合理竞替的方法对国外 10 个信息服务机构的 RDM 政策要素进行了归纳提炼，形成了初步的政策要素，但我国实践过程中可能存在的问题还需要在实践基础上对数据管理政策内容进行细分和修订。

6.3 未来研究展望

随着信息技术的发展，数据资源越来越多，数据管理已成为一种必然趋势。若要激发科研数据的全

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 3 国外信息服务机构 RDM 政策要素

研究问题	政策要素	政策内容提炼
管理什么（科研数据存缴内容规范）	多方面、多角度、多层次支持 RDM	信息服务机构在数据资源进行标引、分类及整理，及确保数据资源的安全性等方面都有非常丰富的经验
	科研数据内容的要求	各类可收集到的数据（如调查数据、实验数据、访谈数据、教学数据、工作文件等）
	科研数据格式的要求	开放的、不依赖于软件的、非专业、未加密、非编译格式（如 ODF、ASCII、MPEG-4、TIFF/JPEG2000、XML 等）
怎么管理（科研数据质量管理规范）	协助创建个人 DMP	科研人员提出数据保存申请时，就需提交 DMP，以保证资源的管理和共享
	积极开展 RDM 培训	开设数据管理课程，帮助图书馆员及科研人员认识科研数据的管理以及科研数据的开放共享
	设计符合需求的元数据标准	不同学科设置不同的元数据标准，既可以帮助科研人员详细了解自己的数据，又可以帮助其他研究人员发现并正确使用及引用数据（如标题、创建者、标识符、日期等）
	规范文件命名格式	文件名要反映文件的特征、内容，要有足够的描述信息以识别数据文件（如项目缩写、实验名称、实验位置等）
	设计合理的科研数据存储时间	受数据类型、资助机构政策、大学政策的影响，数据存储时间一般都在 4 年以上
	建立数据知识库，协助数据存储	机构数据知识库或者其他可以存储并保证数据安全的数据中心等
如何使用（共享规范）	规范科研数据的引用	数据引用必须标明出处，遵守科研道德规范
	协助制定合理的科研数据使用许可及范围	使用许可：参照各图书馆规定 使用范围：科研用户的要求为主，知识库、学校规定为辅
	划分科研数据类型，注重隐私保护	数据一般分为禁止、限制、保密和公开四种类型，根据不同情况采取不同隐私保护措施；当涉及人类受试者数据时，要接受监督并使用专门的数据搜集工具

部潜力，需要进行资助机构政策、研究教育机构政策、学术期刊的数据政策，乃至新兴的数据级别计量（Data Level Metric）等研究与实践。

本文调研了国外现有的信息服务机构的政策，形成了初步框架，未来还需要在访谈科研人员的基础上进一步累积和深化实施方案，包括《最佳实践》和《信息服务机构的数据管理政策指南》等的研究和政策咨询服务。若干研究提问包括但不限于：

（1）如何提高人们对数据引用的意识？激励科研人员真正地去提供他们的科研数据，是否可以设立一些小的、有激励性的榜样赋予科研人员象征性的奖励，例如“每月最佳数据分享者”？

（2）科研数据的管理是为了促使资源能够被更多的科研人员发现，进而产生更多的科研成果，但是科研数据的价值如何能够得到体现？是否只能通过引用？又如何使用指标跟踪数据的引用？除了永久标识符，是否还有其他途径？

（3）如何将 RDM 服务嵌入到大学图书馆的专业科研数据中心？如何建立嵌入到科研院所或大学的基础设施？如何实现并提高 RDM 的嵌入式服务？

（4）科研数据的描述往往是非规范的，是否需要建立术语服务器的受控词表供用户参考？

（5）在我国现有的科研政策体系下，图书馆在数据管理中可以发挥怎样的作用？图书馆员呢？如何针对不同的数据类型及学科特点，制定不同的数据管理政策？

目前中国科学院文献情报中心的科技信息政策中心已经全文翻译了《机构知识库进行科研数据管理的政策指南》一文，而且在 2015 年 10 月预计将举办第四届中国开放获取推介周，期待有更多同行对 RDM 的共同研究与推进。

致谢：中国科学院文献情报中心科技信息政策中心组织编译了《欧盟图书馆实施科研数据管理的案例汇编》报告并在机构知识库（ir. las. ac. cn）上开放共享。

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

参考文献

1 Hey T. Cyberinfrastructure for e-Science[J]. Science, 2005, 308(5723): 817-821

2 科学数据共享工作应该受到尊重和支持——徐冠华院士访问地球系统科学信息共享中心[EB/OL].[2015-02-25]. http://www.most.gov.cn/kjbgz/201002/t20100221_75953.htm

3 吴建中.知识是流动的:出版界与图书馆界的新课题[J].图书馆杂志,2015,34(3):4-11

4 Sykes J. Managing the UK's Research Data: Towards a UK Research Data Service[J]. New Review of Information Networking,2008,14(1):21-36

5 张晓林.开放获取、开放知识、开放创新推动开放知识服务——3O会聚与研究图书馆范式再转变[J].现代图书情报技术,2013,230(2):2-10

6 司莉,邢文明.国外科学数据管理与共享政策调查及对我国的启示[J].情报资料工作,2013(1):61-66

7 崔宇红.E-Science环境中研究图书馆的新角色:科学数据管理[J].图书馆杂志,2012,31(10):20-23

8 Community Capability Model Framework[EB/OL].[2015-03-03]. http://communitymodel.sharepoint.com/Documents/CCMDIRWhitePaper-v1-0.pdf

9 Kim J. Data Sharing and its Implications for Academic Libraries[J]. New Library World,2013,114(11):494-506

10 Shen Y, Varvel V E. Developing Data Management Services at the Johns Hopkins University[J]. Journal of Academic Librarianship, 2013,39(6):552-557

11 Kutay S. Advancing Digital Repository Services for Faculty Primary Research Assets: An Exploratory Study[J]. Journal of Academic Librarianship, 2014,40:642-649

12 Davidson J, Jones S, Molloy L, et al. Emerging Good Practice in Managing Research Data and Research Information within UK Universities[J]. Procedia Computer Science, 2014,33(33): 215-222

13 Marchionini G.科研数据管理保障数据质量促进iSchools新科学研究[J].图书情报知识,2013(4):4-9

14 Wicherts J M, Bakker M. Publish (your data) or (let the data) Perish! Why not publish your data too?[J]. Intelligence,2012, 40(2):73-76

15 顾立平.科学数据权益分析的基本框架[J].图书情报知识,2014(1):34-51

16 吴振新,李丹丹.研究数据管理框架研究[J].图书馆学研究,2012,24:47-52,67

17 吴新年.学术图书馆的科研数据管理服务[J].情报资料工作,2014(5):74-78

18 沈梦轩.国外科学数据资源服务实践研究[J].图书馆学研究,2014,15:89-93

19 黄如花,邱春艳.国内外科学数据元数据研究进展[J].图书与情报,2014(6):102-108

20 杨云秀,顾立平,张瑶,等.国外科研教育机构数据政策的调研与分析——以英国10所高校为例[J].图书情报工作.2015,59(5):53-59

21 张瑶,顾立平,杨云秀,等.国外科研资助机构数据政策的调研与分析——以英美研究理事会为例[J].图书情报工作,2015(6):53-60

22 项英,赖剑菲,丁宁.高校图书馆科学数据管理服务实践探索——以武汉大学社会科学数据管理为例[J].情报理论与实践,2013,36(12):89-93

23 刘霞,饶艳.高校图书馆科学数据管理与服务初探——武汉大学图书馆案例分析[J].图书情报工作,2013,57(6):33-38

24 "CALIS三期特色库子项目验收专家评审会"在我馆顺利召开[EB/OL].[2015-03-03]. http://gzw.lib.whu.edu.cn/pe/Article/ShowArticle.asp?ArticleID=873

25 武汉大学图书馆成功举办"CALIS三期建设项目学科馆员培训教师/学员代表研讨会"[EB/OL].[2015-03-03]. http://gzw.lib.whu.edu.cn/pe/Article/ShowArticle.asp?ArticleID=848

26 丁宁,马浩琴.国外高校科学数据生命周期管理模型比较研究及借鉴[J].图书情报工作,2013,57(6):18-22

27 赖剑菲,洪正国.对高校科学数据管理平台建设的建议[J].图书情报工作,2013,57(6):23-27

28 谢春枝,燕今伟.国内外高校科学数据管理和机制建设研究[J].图书情报工作,2013,57(6):12-17,38

29 胡永生,刘颖.基于用户调查的高校科学数据管理需求分析[J].图书情报工作,2013,57(6):28-32,78

30 洪正国,项英.基于Dspace构建高校科学数据管理平台——以蝎物种与毒素数据库为例[J].图书情报工作,2013,57(6):39-42,84

31 The 9th International Conference on Cooperation and Promotion of Information Resources in Science and Technology (COINFO 201[EB/OL].[2015-03-03]. http://coinfo.istic.ac.cn/index.html#

32 Open Science Case Studies[EB/OL].[2015-03-03]. http://www.rin.ac.uk/our-work/data-management-and-curation/open-science-case-studies

33 Support for Managing Research Data[EB/OL].[2015-03-05]. http://www.lib.cam.ac.uk/dataman/

34 顾立平.前瞻导论[M].台北:设计家出版社,2013

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

35 Tenopir C, Allard S, Douglass K, et al. Data Sharing by Scientists: Practices and Perceptions[J]. PloS one,2011,6(6):1-21
36 Auckland M. Re-Skilling for Research: An Investigation into the Role and Skills of Subject and Liaison Librarians Required to Effectively Support the Evolving Information Needs of Researchers. London: Research Libraries UK. 2012
37 Cox A M, Pinfield S. Research Data Management and Libraries: Current Activities and Future Priorities[J]. Journal of Librarianship and Information Science. 2014,46(4):299-316
38 Gold A K. Cyberinfrastructure, Data, and Libraries, Part 2:Libraries and the Data Challenge: Roles and Actions for Libraries [EB/OL].[2015-03-04]. http://www.dlib.org/dlib/september07/gold/09gold-pt2.html
39 司莉,庄晓喆,王思敏,等. 2005年以来国外科学数据管理与共享研究进展与启示[J].国家图书馆学刊,2013(3):40-49
40 Cox A M, Verbaan E, Sen B. Upskilling Liaison Librarians for Research Data Management[J]. Ariadne,2012,70
41 Martinez-Uribe L, Macdonald S. A New Role for Academic Librarians: Data Curation[EB/OL].[2015-03-04].http://www.era.lib.ed.ac.uk/bitstream/1842/3207/6/data_role.doc
42 Lyon L. The Informatics Transform: Re-Engineering Libraries for the Data Decade[J]. International Journal of Digital Curation. 2012,7(1):126-38
43 科研数据管理与图书馆:当前活动与未来工作重点[EB/OL].[2015-03-04]. http://copyright.las.ac.cn/news/79d178146570636e7ba174064e0e56fe4e669986-5f53524d6d3b52a84e0e672a67655de54f5c91cd70b9
44 Jettens M. Research Data Management at Radboud University [EB/OL].[2015-03-04]. http://libereurope.eu/wp-content/uploads/2014/06/LIBER-Case-Study-Radboud.pdf
45 Overview[EB/OL].[2015-03-04]. http://guides.library.yale.edu/content.php?pid=324929&sid=2665407
46 McNeill K. Research Data Management: 101 The Lifecycle of a Dataset[EB/OL].[2015-03-05].http://libraries.mit.edu/data-management/files/2014/05/research-data-management-iap2014.pdf
47 About Storing & Archiving Your Research Data[EB/OL].[2015-03-05]. http://dmp.data.jhu.edu/preserve-share-research-data/preserve-archive/#Archive
48 What can Phaidra do?[EB/OL].[2015-03-04]. http://phaidraservice.univie.ac.at/en/phaidra/what-can-phaidra-do/
49 File Formats for Long-Term Access[EB/OL].[2015-03-04]. http://libraries.mit.edu/data-management/store/formats/
50 Management of Research Records[EB/OL].[2015-03-05]. www.kcl.ac.uk/aboutkings/governance/compliance/guidance/research/research.pdf
51 Data Preservation at Stanford[EB/OL].[2015-03-05]. http://library.stanford.edu/research/data-management-services/share-and-preserve-research-data/data-preservation-stanford
52 Data Files: Appropriate Formats[EB/OL].[2015-03-05]. http://guides.library.yale.edu/content.php?pid=324929&sid=2665468
53 Write a Data Management Plan[EB/OL].[2015-03-05]. http://libraries.mit.edu/data-management/plan/write/
54 Data Management Plans[EB/OL].[2015-03-04]. http://library.stanford.edu/research/data-management-services/data-management-plans
55 Data Management Planning[EB/OL].[2015-03-04]. http://www.lib.cam.ac.uk/dataman/pages/planning.html
56 Data Management Planning[EB/OL].[2015-03-04]. http://data.research.cornell.edu/content/data-management-planning
57 Writing Data Management or Sharing Plans[EB/OL].[2015-03-04]. http://dmp.data.jhu.edu/assistance/guidance-on-writing-data-management-plans/
58 Data Management Plan[EB/OL].[2015-03-04]. http://www.ru.nl/library/services/research/researchdata/dmp/
59 Data Management Plans[EB/OL].[2015-03-04]. http://guides.library.yale.edu/content.php?pid=324929&sid=2665409
60 Workshops[EB/OL].[2015-03-06]. http://libraries.mit.edu/data-management/services/workshops/
61 Upcoming Events[EB/OL].[2015-03-06]. http://libraries.mit.edu/news/events/
62 Training Resources[EB/OL].[2015-03-06]. http://www.lib.cam.ac.uk/dataman/training.html
63 Data Services[EB/OL].[2015-03-06]. http://guides.nyu.edu/DS_classes
64 JHUDMS Training Sessions[EB/OL].[2015-03-05]. http://dmp.data.jhu.edu/resources/data-management-training-resources/
65 Research Data Management Training[EB/OL].[2015-03-06]. http://www.kcl.ac.uk/library/researchsupport/research-data-management/training.aspx
66 Training[EB/OL].[2015-03-05]. http://www.ru.nl/library/services/research/researchdata/training/
67 Disciplinary Metadata[EB/OL].[2015-03-04]. http://www.dcc.ac.uk/resources/metadata-standards
68 Documentation & Metadata[EB/OL].[2015-03-04]. http://li-

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.   http://www.cnki.net

braries. mit. edu/data-management/store/documentation/
69 Managing Research Data[EB/OL]. [2015-03-05]. http://guides. library. yale. edu/content. php? pid = 324929 & sid = 2665468
70 Data Management Planning[EB/OL]. [2015-03-05]. http://guides. nyu. edu/content. php? pid=394716 & sid=3341859
71 Creating Metadata[EB/OL]. [2015-03-05]. http://library. stanford. edu/research/data-management-services/data-best-practices/creating-metadata
72 Metadata and Describing Data[EB/OL]. [2015-03-05]. http://data. research. cornell. edu/content/writing-metadata
73 Guide to Writing "readme" Style Metadata[EB/OL]. [2015-03-05]. http://data. research. cornell. edu/content/readme
74 Research Data Management: File Organization[EB/OL]. [2015-03-06]. http://159. 226. 251. 229/videoplayer/file-organization-july2014. pdf? ich_u_r_i=3f7baa2b82f7161a39fdf7031351c15e & ich_s_t_a_r_t = 0 & ich_e_n_d = 0 & ich_k_e_y = 1545048911750563432409 & ich_t_y_p_e=1 & ich_d_i_s_k_i_d=2 & ich_u_n_i_t=1
75 File Management[EB/OL]. [2015-03-05]. http://data. research. cornell. edu/content/file-management
76 Best Practices for File Naming[EB/OL]. [2015-03-05]. http://library. stanford. edu/research/data-management-services/data-best-practices/best-practices-file-naming
77 How Long Should I Keep My Research Data? [EB/OL]. [2015-03-08]. http://www. kcl. ac. uk/aboutkings/governance/compliance/guidance/research/decision. pdf
78 What can Phaidra do? [EB/OL]. [2015-03-08]. http://phaidraservice. univie. ac. at/en/phaidra/what-can-phaidra-do/
79 Johns Hopkins Data Archive Dataverse[EB/OL]. [2015-03-08]. Network https://archive. data. jhu. edu/dvn/
80 DSpace[EB/OL]. [2015-03-06]. https://www. repository. cam. ac. uk/
81 u:scholar as a Institutional Repository of the University of Vienna [EB/OL]. [2015-03-08]. http://phaidraservice. univie. ac. at/en/to-use-phaidra-efficiently/uscholar-as-a-institutional-repository-of-the-university-of-vienna/
82 Data Deposit[EB/OL]. [2015-03-05]. https://lib. stanford. edu/sdr/data-deposit
83 Storing Data[EB/OL]. [2015-03-05]. http://www. lib. cam. ac. uk/dataman/pages/storage. html
84 Find a Data Repository[EB/OL]. [2015-03-05]. http://libraries. mit. edu/data-management/share/find-repository/
85 **邢文明,吴方枝,司莉. 高校图书馆开展科研数据管理与共享服务调查分析**[J]. **图书馆论坛**,2013,33(6):19-25,61
86 Data Repositories[EB/OL]. [2015-03-05]. http://guides. nyu. edu/content. php? pid=394716 & sid=3341905
87 **顾立平. 数据级别计量-概念辨析与实践进展**[J]. **中国图书馆学报**,2015, 216(41):56-71
88 Social Science Data Services: Cite data[EB/OL]. [2015-03-05]. http://libguides. mit. edu/ssds/cite
89 Data Citation[EB/OL]. [2015-03-05]. http://dmp. data. jhu. edu/preserve-share-research-data/sharing-your-research-data/
90 Data sharing & re-use[EB/OL]. [2015-03-05]. http://guides. library. yale. edu/content. php? pid=324929 & sid=2665439
91 Citation Management[EB/OL]. [2015-03-06]. https://www. library. cornell. edu/research/citation
92 Citing Data and Getting Cited[EB/OL]. [2015-03-05]. http://www. lib. cam. ac. uk/dataman/pages/citations. html
93 **张晓林,张冬荣,李麟,等. 机构知识库内容保存与传播的权利管理**[J]. **中国图书馆学报**,2012(4):46-54
94 **张晓林,张冬荣,李麟,等. 公共教育科研单位机构知识库内容存缴与传播的权益管理政策指南**[EB/OL]. [2015-03-08]. http://ir. las. ac. cn/handle/12502/5043
95 Looking after your dData[EB/OL]. [2015-03-08]. http://www. lib. cam. ac. uk/dataman/lookafter. html
96 Licensing[EB/OL]. [2015-03-08]. http://library. stanford. edu/research/data-management-services/ share-and-preserve-research-data/licensing
97 Confidentiality and Intellectual Property[EB/OL]. [2015-03-08]. http://libraries. mit. edu/data-management/share/confidentiality/
98 Storage and Backup of Sensitive Data[EB/OL]. [2015-03-08]. http://library. stanford. edu/research/data-management-services/storage-and-backup/sensitive-data
99 Sharing Sensitive[EB/OL]. [2015-03-08]. datahttp://library. stanford. edu/research/data - management-services/share-and-preserve-research-data/sharing-sensitive-data
100 Working with Sensitive Data[EB/OL]. [2015-03-08]. http://library. stanford. edu/research/data - management-services/data-best-practices/working-sensitive-data
101 Data Protection (Personal Information Management) Policy [EB/OL]. [2015-03-08]. http://www. kcl. ac. uk/college/policyzone/assets/files/governance_and_legal/DataProtectionPolicy_updated_Oct2011. pdf

(**收稿日期**:2015-05-20)

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net